The Annals of Probability
1993, Vol. 21, No. 2, 800–818

ASYMPTOTIC EXPANSIONS FOR MARTINGALES[1]

By Per Aslak Mykland

University of Chicago

The paper contains a "smoothed" one-step triangular array asymptotic expansion for discrete-time martingales. An important element of the proof is a second-order description of Skorokhod embedding of discrete martingales in continuous ones. An application to Markov processes is given, along with a bootstrapping example.

1. Introduction. The use of Edgeworth expansions in statistical inference has a long history, going back to Edgeworth (1883, 1905), Cramér (1937) and Cornish and Fisher (1937). This approach has recently received renewed attention, both because of new breakthroughs in research on expansions [see, e.g., Bhattacharya and Rao (1976, 1986), Bhattacharya and Ghosh (1978), Barndorff-Nielsen and Cox (1984), McCullagh (1984a, b) and Abramovitch and Singh (1985)], and also because of its relevance to bootstrapping [see, e.g., Singh (1981), Hall (1986, 1988), Beran (1987, 1988) and Efron (1987)]. Though traditionally confined to inference situations with independent observations, asymptotic expansions have recently been developed also for inference in dependent variable models. Research has been conducted using assumptions of weak dependence [Goetze and Hipp (1983), Bose (1986, 1987, 1988) and Jensen (1986)], Markov dependence [Malinovskii (1987) and Jensen (1989)], ARMA structure [Taniguchi (1984) and Tanaka (1986)] and martingale structure [Mykland (1992)]. The approach of Barndorff-Nielsen and Cox (1984) and McCullagh (1984a) also lends itself to studying dependent variables.

The purpose of this paper is to develop an Edgeworth expansion for general scalar martingales, thus extending our previous work on the subject [Mykland (1989, 1992)]. Our reason for focusing on martingales is that they seem particularly well suited to inference, in particular in parametric models, in time series and in survival analysis. Good illustrations of this are Klimko and Nelson (1978), Hall and Heyde (1980), Chapter 6, Tjøstheim (1986), Andersen and Gill (1982), Andersen and Borgan (1985) and Wong (1986).

Apart from the expansion theorem itself (Section 2), the paper gives two applications of the result: the bootstrapping of an AR(1) process (Section 3) and expansions for Markov sums (Section 4).

Received August 1990; revised September 1991.

[1]Research supported in part by NSF Grant DMS-89-02667. This manuscript was prepared using computer facilities supported in part by NSF Grants DMS-89-05292, DMS-87-03942 and DMS-86-01732 awarded to the Department of Statistics at the University of Chicago, and by the University of Chicago Block Fund.

AMS 1991 *subject classifications*. Primary 60F99, 60G42, 62E20; secondary 60J05, 62M09, 62M10.

Key words and phrases. Edgeworth expansions, bootstrapping, Markov processes, martingales, Skorokhod embedding, time series.

The proofs are all in Section 5. A main methodology is a second-order description of Skorokhod embedding of a discrete-time martingale in a continuous one; see (5.1) and Lemma 1 (in Section 5). A consequence of this characterization is that distance from normality is best described by a linear combination of observed and predictable square variation (with weights of 1/3 and 2/3, respectively)—see (2.16).

2. Expansions for discrete-time martingales. We shall be dealing with a triangular array $(l_t^{(N)},\ t = 0, 1, 2, \dots, T_N)$, $N = 1, 2, \dots,$ of zero-mean martingales, say

$$l_t^{(N)} = X_1^{(N)} + X_2^{(N)} + \cdots + X_t^{(N)}, \tag{2.1}$$

where the $X_i^{(N)}$'s are martingale differences. For many purposes it is also useful to studentize the martingale, so that we study the asymptotic behavior of $l_{T_N}^{(N)}/\sqrt{\sigma_N}$, where σ_N is some (typically random) normalizing factor.

The expansion result is for the distribution function

$$F_N(x) = P\left(\frac{l_{T_N}^{(N)}}{\sqrt{\sigma_N}} \le x \,\middle|\, \sigma_N > 0 \right). \tag{2.2}$$

The expansion is of the form

$$\int_{-\infty}^{+\infty} g(x)\, dF_N(x) = \int_{-\infty}^{+\infty} g(x)\, d\Phi(\beta^{-1}x) + r_N J(g) + o(r_N), \tag{2.3}$$

and holds uniformly over large classes of twice differentiable functions g. For exact statements we refer to Theorem 1.

The rate of the expansion, r_N, need not be the same as $N^{-1/2}$, though this will often be the case. The rate is determined by conditions (2.6)–(2.14) below. The expansion term, $J(g)$, also need not have the standard form. It is defined below, in (2.15) and (2.18)–(2.19).

The expansion does, in general, *not hold* when g is an indicator function of an interval. Ours is a "smoothed" expansion. There is no Cramér-type condition, and in return the expansion does not hold in a pointwise topology. Similar results for i.i.d. variables have been developed by Goetze and Hipp (1978). It is not clear what a Cramér-type condition would look like: The literature on martingale Berry–Esseen bounds [see, e.g., Bolthausen (1982) and Haeusler (1988)] bears witness to that. Already with the weak dependence conditions of Goetze and Hipp (1983)—which are, on the whole, much stronger than the conditions of this paper—the Cramér condition is quite complex.

We now turn to the conditions needed for the result. These will be imposed on σ_N and on variation measures associated with martingales, the optional and predictable kth-order variations, denoted respectively by $[l^{(N)}, \dots, l^{(N)}]$ and $\langle l^{(N)}, \dots, l^{(N)} \rangle$. They are defined by

$$\underbrace{[l^{(N)}, \dots, l^{(N)}]_t}_{k \text{ times}} = \sum_{i=1}^{t} \left(X_i^{(N)}\right)^k \tag{2.4}$$

and, whenever it is defined,

$$\underbrace{\langle l^{(N)}, \dots, l^{(N)} \rangle_t}_{k \text{ times}} = \sum_{i=1}^{t} E\Big(\big(X_i^{(N)}\big)^k \big| \mathscr{F}_{i-1}^{(N)}\Big), \tag{2.5}$$

where $(\mathscr{F}_t^{(N)})$ is the "history" or "filtration"; cf. Brémaud (1981), Chapter 1.1.

The basic assumptions on σ_N and the variation measures are integrability and central limit conditions. In addition to r_N, we need the nonrandom sequence $c_N, c_N^{1/2}$ being $O_p(l_{T_N}^{(N)})$. One can without loss of generality let $c_N = 1$, by assimilating it into $l_{T_N}^{(N)}$. We now state the conditions.

Integrability condition for the fourth-order variation.

$$E[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N} = O(c_N^2 r_N^2). \tag{2.6}$$

Integrability conditions for the square variation. There are constants b^2, $\bar{k}$ and $\underline{k}$ so that $\infty \geq \bar{k} > b^2 > \underline{k} \geq 0$ and so that, for $(l^{(N)}, l^{(N)})_{T_N}$ being either $[l^{(N)}, l^{(N)}]_{T_N}$ or $\langle l^{(N)}, l^{(N)} \rangle_{T_N}$,

$$r_N^{-1}\left(\frac{(l^{(N)}, l^{(N)})_{T_N}}{c_N} - b^2\right) I\left\{\bar{k} \geq \frac{(l^{(N)}, l^{(N)})_{T_N}}{c_N} \geq \underline{k}\right\} \tag{2.7}$$

is uniformly integrable,

I being the indicator function, and so that

$$P\left(\bar{k} \geq \frac{(l^{(N)}, l^{(N)})_{T_N}}{c_N} \geq \underline{k}\right) = 1 - o(r_N) \tag{2.8}$$

(it is shown in Consequence 1 in Section 5 that the condition on $[l^{(N)}, l^{(N)}]_{T_N}$ is equivalent to the condition on $\langle l^{(N)}, l^{(N)} \rangle_{T_N}$, provided the above assumption on $[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N}$ holds).

Integrability conditions for σ_N. There are measurable sets D_T^* and constants b_*^2 and δ ($\delta > 0$) so that

$$\sup_N E\left[\frac{1}{r_N}\left|\frac{\sigma_N}{c_N} - b_*^2\right|\right]^{1+\delta} I_{D_T^*} < \infty \tag{2.9}$$

and so that

$$P(D_T^*) = 1 - o(r_N). \tag{2.10}$$

The central limit condition. There are Borel-measurable functions ψ_o, ψ_p and ψ_*, so that, whenever

$$\begin{gathered}\left(b^{-1}\frac{l_{T_N}^{(N)}}{\sqrt{c_N}}, r_N^{-1}\left(\frac{[l^{(N)}, l^{(N)}]_{T_N}}{c_N} - b^2\right),\right.\\ \left. r_N^{-1}\left(\frac{\langle l^{(N)}, l^{(N)}\rangle_{T_N}}{c_N} - b^2\right), r_N^{-1}\left(\frac{\sigma_N}{c_N} - b_*^2\right)\right) \qquad (2.11)\\ \xrightarrow{\mathscr{L}} (Z, \xi_o, \xi_p, \xi_*)\end{gathered}$$

as $N \to \infty$ through a subset of the integers, then

$$E(\xi_o|Z) = b^2\psi_o(Z) \quad \text{a.s.,} \tag{2.12}$$

$$E(\xi_p|Z) = b^2\psi_p(Z) \quad \text{a.s.,} \tag{2.13}$$

and

$$E(\xi_*|Z) = b_*^2\psi_*(Z) \quad \text{a.s.} \tag{2.14}$$

This rounds off our list of conditions.

The central limit condition gives us the quantities needed to define the expansion term: For all functions g for which this is defined, set

$$J(g) = \tfrac{1}{2}E\left[\beta^2\psi(Z)g''(\beta Z) - \psi_*(Z)\beta Z g'(\beta Z)\right], \tag{2.15}$$

where

$$\psi = \tfrac{1}{3}\psi_o + \tfrac{2}{3}\psi_p \tag{2.16}$$

and where β is the asymptotic standard deviation of $l_{T_N}^{(N)}/\sqrt{\sigma_N}$, that is,

$$\beta = bb_*^{-1}. \tag{2.17}$$

Our conditions imply those for the martingale CLT [see, e.g., Hall and Heyde (1980), page 58, Theorem 3.2 and Corollary 3.1), so Z is standard normal. In general, we do not know anything about the distributions of ξ_o, ξ_p and ξ_*, and they need not be the same for different subsequences. In practice, (Z, ξ_o, ξ_p, ξ_*) is often Gaussian, but this need not be so; cf. Section 2.3 in Mykland (1992). Also, the ξ's can be degenerate: ξ_* if σ_N is nonrandom; ξ_p if the X_t's are independent.

We are now in a position to state the expansion theorem.

THEOREM 1. *Suppose that the triangular array $(l_t^{(N)})_{0 \le t \le T_N}$, $N = 1, 2, \dots,$ and the normalizing factor σ_N satisfy the integrability and central limit conditions above, with $r_N = o(1)$. Then (2.3) holds, the convergence being uniform on sets $\mathscr{C}$ on functions g which are twice differentiable, with g, g′ and g″ uniformly bounded, and with $\{g'', g \in \mathscr{C}\}$ being equicontinuous a.e. Lebesgue.*

There is a certain structural similarity between our expansion theorem and the martingale central limit theorem. In both cases, the results state that one level of asymptotics gives you the next level for free: The main condition for the CLT is a law of large numbers for either $[l^{(N)}, l^{(N)}]_{T_N}/c_N$ or $\langle l^{(N)}, l^{(N)}\rangle_{T_N}/c_N$; the main condition for the expansion is a CLT for either of the same quantities.

Subject to some minimum niceness on the part of ψ and ψ_*, integration by parts yields that

$$J(g) = \frac{1}{2}\int_{-\infty}^{+\infty} g(x)\, d\big(\lambda(\beta^{-1}x)\phi(\beta^{-1}x)\big), \tag{2.18}$$

where

$$\lambda(x) = \psi'(x) - \psi(x)x + \psi_*(x)x. \tag{2.19}$$

If we denote $o_2(r_N)$ to describe the kind of convergence used in the theorem [cf. Mykland (1992), Remark 2.3], the result in Theorem 1 is that

$$F_N(x) = \Phi(\beta^{-1}x) + r_N\tfrac{1}{2}\lambda(\beta^{-1}x)\phi(\beta^{-1}x) + o_2(r_N), \tag{2.20}$$

which is a more standard way of stating an expansion.

When it comes to verifying the conditions of Theorem 1 in practice, it is worth noting that $[l^{(N)}, l^{(N)}]_t - \langle l^{(N)}, l^{(N)}\rangle_t$ is a martingale. Under conditions which follow easily from Theorem 3.2 and Corollary 3.1 of Hall and Heyde [(1980), page 58] the joint limiting distribution of $([l^{(N)}, l^{(N)}]_{T_N} - \langle l^{(N)}, l^{(N)}\rangle_{T_N})/r_N c_N$ and $l^{(N)}_{T_N}/\sqrt{c_N}$ can be described, whence $\psi_o - \psi_p$ can be obtained. In fact,

$$(\psi_o - \psi_p)(z) = zb^{-3} \lim_{N\to\infty} r_N^{-1} c_N^{-3/2} \langle l^{(N)}, l^{(N)}, l^{(N)}\rangle_{T_N}, \tag{2.21}$$

"lim" meaning limit in probability.

If we are dealing with a nontriangular array, weaker conditions for this to occur can be stated.

PROPOSITION 2. *Let $l_t^{(N)}$ be the same for all N. Assume that the integrability conditions for the fourth-order and square variations are satisfied, and that the right-hand side of* (2.21) *is well defined and nonrandom. Then ψ_o is well defined if and only if ψ_p also is, and, if so,* (2.21) *holds.*

The formula (2.21) is related to a third-order cumulant. Indeed, if the X's are independent, ψ_p is 0, and $(\psi_o - \psi_p)/3$ is the only term in ψ which remains.

3. A bootstrapping example. Being also valid for triangular arrays, the martingale expansion lends itself to the study of the second-order properties of bootstrapping. As a very simple example, we shall discuss the AR(1) process. We shall see that, in the $o_2(N^{-1/2})$ sense, the second-order correctness of the bootstrap holds under very weak assumptions. Conditions for the second-order correctness in the $o(N^{-1/2})$ sense have been obtained by Bose (1988), and are considerably stronger.

What is needed is a triangular array result for the distribution function

$$F_N(x;\theta,Q) = P_{\theta,Q}\Big(s_N^{-1/2}\big(\hat{\theta}_N - \theta\big) \le x\Big), \tag{3.1}$$

where the process is given by

$$\begin{aligned} \eta_{t+1} &= \theta\eta_t + \varepsilon_{t+1}, \qquad 0 \le t \le N-1, \\ \eta_0 &= x_0, \end{aligned} \tag{3.2}$$

the ε's being i.i.d. with distribution Q. $\hat{\theta}_N$ is the conditional least squares estimator, $\hat{\theta}_N = \sum_{t=0}^{N-1}\eta_{t+1}\eta_t / \sum_{t=0}^{N-1}\eta_t^2$, and $s_N = (1-\hat{\theta}_N^2)/N$. The result is as follows.

PROPOSITION 3. *Suppose that the nonrandom sequences θ_N and Q_N satisfy $E_{Q_N}\varepsilon = 0$,*

$$E_{Q_N}\varepsilon^2 = \kappa_2 + O(N^{-1/2}), \tag{3.3}$$

$$E_{Q_N}\varepsilon^3 = \kappa_3 + o(1), \tag{3.4}$$

$$E_{Q_N}\varepsilon^4 = O(1) \tag{3.5}$$

and

$$\theta_N = \theta + O(N^{-1/2}), \tag{3.6}$$

where $|\theta| < 1$. Suppose that x_0 is fixed across N. Then

$$\begin{aligned} &F_N(x;\theta_N,Q_N) \\ &\quad = \Phi(x) + N^{-1/2}\phi(x)\left\{\frac{1}{6}\frac{\kappa_3^2}{\kappa_2^3}\frac{\left(1-\theta^2\right)^{3/2}}{1-\theta^3}(1-x^2)\right. \\ &\qquad\qquad \left. + \frac{\theta}{\left(1-\theta^2\right)^{1/2}}\right\} + o_2(N^{-1/2}). \end{aligned} \tag{3.7}$$

A bootstrap estimate of the distribution function $F_N(x;\theta,Q)$ is $F_N(x;\hat{\theta}_N,\hat{Q}_N)$, where $\hat{Q}_N$ is the empirical of the residuals, centered to have mean 0. The following corollary is then obvious.

COROLLARY 4. *Suppose that $E_Q\varepsilon^4 < \infty$, and that $|\theta| < 1$. Then*

$$F_N\big(x;\hat{\theta}_N,\hat{Q}_N\big) = F_N(x;\theta,Q) + o_2(N^{-1/2}) \quad \textit{in probability}. \tag{3.8}$$

4. Application to Markov chains. Let $\eta_0, \eta_1, \dots, \eta_t, \dots$ be a Harris ergodic Markov chain [cf. Nummelin (1984), Chapter 6.3, page 114], with initial distribution λ and stationary distribution π. Let f be a function satisfying that, for all t,

$$E\big(f(\eta_{t+1})|\eta_t\big) = 0 \quad \text{a.s.}, \tag{4.1}$$

so that

$$l_N = \sum_{t=0}^{N} f(\eta_t) \tag{4.2}$$

is a martingale. Now suppose that

$$E_\pi f(\eta_0)^4 < \infty. \tag{4.3}$$

Let s_B be the hitting time for the set B. Suppose that, for all measurable B, $\pi(B) > 0$,

$$E_\pi s_B < \infty, \tag{4.4}$$

$$E_\lambda s_B < \infty, \tag{4.5}$$

$$E_\pi \sum_{t=1}^{s_B} f(\eta_t)^2 < \infty, \tag{4.6}$$

$$E_\pi f(\eta_0)^2 \sum_{t=1}^{s_B} f(\eta_t)^2 < \infty \tag{4.7}$$

and

$$E_\lambda \sum_{t=0}^{s_B} |f(\eta_t)| < \infty. \tag{4.8}$$

In usual Markov terminology, this is to say that λ must be 1- and $|f|$-regular, π must be 1- and f^2-regular, and $f^2(x)\pi(dx)$ must be f^2-regular [Nummelin (1984), Definition 5.4, page 79].

If $\lambda = \pi$ and f is bounded, this reduces to π having to be ergodic of degree 2 [Nummelin (1984), Chapters 5.4 and 6.4]. If $\lambda = \pi$ and π is geometrically ergodic [Nummelin (1984), Chapter 5.5], then it is enough if $E_\pi|f(\eta)|^{4+\delta} < \infty$ for any $\delta > 0$. Many processes are geometrically ergodic; see Chan (1990) and Tong (1990) for a wealth of examples. Intermediate types of ergodicity require intermediate moment conditions.

With these conditions, we get the following result.

PROPOSITION 5. *Let conditions* (4.1)–(4.8) *above be satisfied. Let* σ_N *satisfy the integrability condition of Section* 2 [(2.9) *and* (2.10)], *and assume that* ψ_* *exists. Then the conditions of Theorem* 1 *are satisfied, and*

$$b^2 = E_\pi f(\eta)^2, \tag{4.9}$$

$$\psi_o(z) = zb^{-3}E_\pi f(\eta_o) \sum_{t=0}^{\infty} \left[f(\eta_t)^2 - E_\pi f(\eta)^2 \right] \tag{4.10}$$

and

$$\psi_p(z) = zb^{-3}E_\pi f(\eta_o) \sum_{t=1}^{\infty} \left[f(\eta_t)^2 - E_\pi f(\eta)^2 \right]. \tag{4.11}$$

5. Proofs. We shall throughout this section assume that $c_N = 1$. This is without loss of generality.

PROOF OF THEOREM 1. The strategy of the proof is as follows. Continuous martingales $\bar{l}_t^{(N)}$, $0 \le t \le T_N$, are created, satisfying $\bar{l}_i^{(N)} = l_i^{(N)}$ for all integers i. The theorem is then proved by verifying the conditions of Theorem 2.2 of Mykland (1992), which deals with asymptotic expansions for continuous martingales. To do this, one needs to verify integrability and central limit conditions for $\langle \bar{l}^{(N)}, \bar{l}^{(N)} \rangle_{T_N}$. These can be derived from the similar conditions on $[l^{(N)}, l^{(N)}]_{T_N}$ and $\langle l^{(N)}, l^{(N)} \rangle_{T_N}$ by controlling the behavior of the (martingale) remainder term

$$m_i^{(N)} = \langle \bar{l}^{(N)}, \bar{l}^{(N)} \rangle_i - \tfrac{1}{3}[l^{(N)}, l^{(N)}]_i - \tfrac{2}{3}\langle l^{(N)}, l^{(N)} \rangle_i. \tag{5.1}$$

Specifically, $(\bar{l}_t^{(N)})$ is created by using Theorem 1 of Heath (1977), which is a form of Skorokhod embedding. It is clear from the statement and proof of that theorem that $(\bar{l}_t^{(N)})$ has a filtration $(\overline{\mathscr{F}}_t^{(N)})$ satisfying $\mathscr{F}_t^{(N)} \subseteq \overline{\mathscr{F}}_t^{(N)}$, and also that $\operatorname{var}(\bar{l}_i^{(N)} | \overline{\mathscr{F}}_{i-1}^{(N)}) = \operatorname{var}(l_i | \mathscr{F}_{i-1}^{(N)})$.

To control the martingale from (5.1), we need the two following lemmas. The first is the technical core of the result, and it will be shown immediately following this proof. Establishing the second lemma is an elementary exercise in inequalities.

LEMMA 1 (Second-order properties of Skorokhod embedding). *Under the assumptions of Theorem* 1,

$$E\left(m_{T_N}^{(N)}\right)^2 = O\left(r_N^2\right), \tag{5.2}$$

$(r_N^{-1} m_{T_N}^{(N)}, l_{T_N}^{(N)})$ *is tight, and whenever it converges in law to, say,* (m, l), *then* $E(m|l) = 0$ *a.s.*

LEMMA 2. *Let* a_N *and* a'_N *be random variables, and let the constant* b^2 *be given. Let*

$$E(a_N - a'_N)^2 = O\left(r_N^2\right). \tag{5.3}$$

Then the statement (S) "*there are constants* $\bar{k}$, b *and* $\underline{k}$, $\infty \ge \bar{k} > b^2 > \underline{k} \ge 0$, *such that*

$$r_N^{-1} a_N I\left(\bar{k} \ge a_N - b^2 \ge \underline{k}\right) \text{ is uniformly integrable} \tag{5.4}$$

and

$$P\left(\bar{k} \ge a_N - b^2 \ge \underline{k}\right) = 1 - o(r_N)," \tag{5.5}$$

is equivalent to the same statement for a'_N (*for some choice of* $\bar{k}$, $\underline{k}$).

The second lemma has the following implications:

CONSEQUENCE 1. Let $a_N = [l^{(N)}, l^{(N)}]_{T_N}$ and $a'_N = \langle l^{(N)}, l^{(N)} \rangle_{T_N}$. Now

$$
\begin{aligned}
E(a_N - a'_N)^2 &= E\left(\sum_{i=1}^{T_N} X_i^2 - E(X_i^2 | \mathscr{F}_{i-1}) \right)^2 \\
&= \sum_{i=1}^{T_N} E\big(\operatorname{Var}(X_i^2 | \mathscr{F}_{i-1})\big) (\text{by the martingale property}) \qquad (5.6)\\
&\leq E[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N} \\
&= O(r_N^2),
\end{aligned}
$$

hence the equivalence between the integrability conditions on $[l^{(N)}, l^{(N)}]_{T_N}$ and $\langle l^{(N)}, l^{(N)} \rangle_{T_N}$ in the presence of the integrability condition on $[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N}$.

CONSEQUENCE 2. Let $a_N = [l^{(N)}, l^{(N)}]_{T_N}$ or $= \langle l^{(N)}, l^{(N)} \rangle_{T_N}$, and let

$$
a'_N = \tfrac{1}{3}[l^{(N)}, l^{(N)}]_{T_N} + \tfrac{2}{3}\langle l^{(N)}, l^{(N)} \rangle_{T_N}. \qquad (5.7)
$$

By exactly the same reasoning as in Consequence 1, a'_N now satisfies the statement (S) of Lemma 2.

CONSEQUENCE 3. Let a'_N be as in Consequence 2, and let $a_N = \langle \bar{l}^{(N)}, \bar{l}^{(N)} \rangle_{T_N}$. Since $E(m_{T_N}^{(N)})^2 = O(r_N^2)$, a_N satisfies the statement (S) of Lemma 2.

If we combine this last fact with the result in Lemma 1 that $E(m|l) = 0$, our Theorem 1 follows from Theorem 2.2 of Mykland (1992).

PROOF OF LEMMA 1. First, use Itô's formula [see Jacod and Shiryaev (1987), Theorem I.4.57, page 57] to differentiate $(\bar{l}_t^{(N)} - \bar{l}_i^{(N)})^2$ for $t \geq i$. This yields

$$
\begin{aligned}
&\left(\bar{l}_{i+1}^{(N)} - \bar{l}_i^{(N)}\right)^2 - \left(\langle \bar{l}^{(N)}, \bar{l}^{(N)} \rangle_{i+1} - \langle \bar{l}^{(N)}, \bar{l}^{(N)} \rangle_i\right) \\
&\qquad = 2\int_i^{i+1} \left(\bar{l}_s^{(N)} - \bar{l}_i^{(N)}\right) d\bar{l}_s^{(N)}.
\end{aligned} \qquad (5.8)
$$

Hence

$$
\begin{aligned}
&E\left\{ \left[\left(\bar{l}_{i+1}^{(N)} - \bar{l}_i^{(N)}\right)^2 - \left(\langle \bar{l}^{(N)} \rangle_{i+1} - \langle \bar{l}^{(N)} \rangle_i\right) \right]^2 \Big| \mathscr{F}_i^{(N)} \right\} \\
&\qquad = 4E\left\{ \int_i^{i+1} \left(\bar{l}_s^{(N)} - \bar{l}_i^{(N)}\right)^2 d\langle \bar{l}^{(N)} \rangle_s \Big| \mathscr{F}_i^{(N)} \right\} \qquad (5.9)\\
&\qquad = \tfrac{2}{3} E\left\{ \left(\bar{l}_{i+1}^{(N)} - \bar{l}_i^{(N)}\right)^4 \Big| \mathscr{F}_i^{(N)} \right\},
\end{aligned}
$$

again by Itô's formula. Note that integrability follows from right to left, using, if necessary, that a continuous martingale can be stopped for any value of its quadratic variation. As

$$(5.10)\qquad \begin{aligned}\Delta m_i^{(N)^2} \le 2\Big(\Delta\big\langle \bar{l}^{(N)}, \bar{l}^{(N)}\big\rangle_i - \Delta[l^{(N)}, l^{(N)}]_i\Big)^2 \\ + \tfrac{8}{9}\big(\Delta[l^{(N)}, l^{(N)}]_i - \Delta\langle l^{(N)}, l^{(N)}\rangle_i\big)^2,\end{aligned}$$

it follows that

$$(5.11)\qquad \begin{aligned} Em_i^{(N)^2} &\le \tfrac{4}{3}\sum_{j=0}^{i-1} E\Big\{\big(l_{j+1} - l_j\big)^4\Big\} \\ &\quad + \tfrac{8}{9}\sum_{j=0}^{i-1} E\Big\{\big[X_j^2 - E\big(X_j^2|\mathscr{F}_{j-1}\big)\big]^2\Big\} \\ &\le \tfrac{20}{9}E[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_i. \end{aligned}$$

By the integrability condition (2.6) on $[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N}$, (5.2) follows.

We now turn to the behavior of $m_i^{(N)} l_i^{(N)}$. (5.8) yields that

$$(5.12)\qquad \begin{aligned} &E\Big\{\Big[\big(\bar{l}_{i+1}^{(N)} - \bar{l}_i^{(N)}\big)^2 - \big(\langle \bar{l}^{(N)}, \bar{l}^{(N)}\rangle_{i+1} - \langle \bar{l}^{(N)}, \bar{l}^{(N)}\rangle_i\big)\Big]\big(\bar{l}_{i+1}^{(N)} - \bar{l}_i^{(N)}\big)\Big|\mathscr{F}_i\Big\} \\ &= E\Big\{2\int_i^{i+1}\big(\bar{l}_s^{(N)} - \bar{l}_i^{(N)}\big)\, d\big\langle \bar{l}^{(N)}, \bar{l}^{(N)}\big\rangle_s \Big|\mathscr{F}_i\Big\} \\ &= \tfrac{2}{3}E\Big\{\big(\bar{l}_{i+1}^{(N)} - \bar{l}_i^{(N)}\big)^3\Big|\mathscr{F}_i\Big\}, \end{aligned}$$

which is well defined since $E|X_{i+1}^{(N)}|^3 < \infty$, and with integrability following from right to left as in (5.9). On the other hand,

$$(5.13)\qquad E\Big[\Big[X_{i+1}^{(N)^2} - E\big(X_{i+1}^{(N)^2}\big|\mathscr{F}_i\big)\Big]X_{i+1}^{(N)}\Big|\mathscr{F}_i\Big] = E\big(X_{i+1}^{(N)^3}\big|\mathscr{F}_i\big).$$

Putting the two previous expressions together shows that $m_i^{(N)} l_i^{(N)}$ is an $(\mathscr{F}_i)$-martingale.

From (5.2), $r_N^{-1} m_{T_N}^{(N)}$ is a tight sequence, and the remarks following the statement of Theorem 1 show that $l_{T_N}^{(N)}$ is tight, whence the pair is tight. We shall show that for all convergent-in-law subsequences, $(l_{T_N}^{(N)}, r_N^{-1} m_{T_N}^{(N)}) \to (l, m)$, $E(m|l) = 0$ a.s.

Assume in the following that

$$(5.14)\qquad \langle l^{(N)}, l^{(N)}\rangle_{T_N} \le \bar{k}.$$

This can be done without loss of generality since we can replace $l_i^{(N)}$ and $m_i^{(N)}$ by $l_{i\wedge t_N}^{(N)}$ and $m_{i\wedge t_N}^{(N)}$, where

$$(5.15)\qquad t_N = \inf\big\{i : \langle l^{(N)}, l^{(N)}\rangle_{i+1} > \bar{k}\big\}$$

or $= T_N$ if the above set is empty. This is a stopping time by Chapter X-1a of Jacod (1979). All our conditions on $l_i^{(N)}$ are satisfied for $l_{i\wedge t_N}^{(N)}$, and the limits for $l_{t_N}^{(N)}$ and $m_{t_N}^{(N)}$ are as for $l_{T_N}^{(N)}$ and $m_{T_N}^{(N)}$, since

$$P(t_N < T_N) = P\big(\langle l^{(N)}, l^{(N)}\rangle_{T_N} > \bar{k}\big) \to 0. \tag{5.16}$$

Also in the following, we shall consider $l^{(N)}$ and $m^{(N)}$ to be embedded in continuous time martingales with right-continuous sample paths, which jump at integer times and are otherwise constant. We shall still use the notation $l^{(N)}$ and $m^{(N)}$. Now set

$$A_N(t) = \begin{cases} \inf\{s: \langle l^{(N)}, l^{(N)}\rangle_s > t\}, \\ T_N, \quad \text{if the above set is empty.} \end{cases} \tag{5.17}$$

As with (5.15), $A_N(t)$ is a family of stopping times. Define $\tilde{l}_t^{(N)} = l_{A(t)}^{(N)}$, and similarly for $\tilde{m}_t^{(N)}$.

Note that, for $t \le \langle l^{(N)}, l^{(N)}\rangle_{T_N}$,

$$t \le \big\langle \tilde{l}^{(N)}, \tilde{l}^{(N)}\big\rangle_t \le t + \Delta\langle l^{(N)}, l^{(N)}\rangle_{A(t)} \quad \text{a.s.,} \tag{5.18}$$

and that, by Jensen's inequality

$$\begin{aligned} &P\Big(\sup_t \big|\Delta\langle l^{(N)}, l^{(N)}\rangle_{A(t)}\big| > \varepsilon\Big) \\ &\quad \le \text{constant} \times E[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N} \\ &\quad \to 0. \end{aligned} \tag{5.19}$$

Using Theorem VI.2.15 (pages 306 and 307) and the discussion in Chapter VI.3b.2 (pages 316 and 317) of Jacod and Shiryaev (1987), it follows that $\langle \tilde{l}^{(N)}, \tilde{l}^{(N)}\rangle_t$ converges in law to $t \wedge b^2$ in the space $\mathbb{D}(\mathbb{R}^+)$ (with the Skorokhod topology). Jacod and Shiryaev, Theorem VI.4.13, page 322, then yields that $\tilde{l}_t^{(N)}$ is tight (with respect to the same weak convergence). Since

$$\begin{aligned} &P\Big(\sup_i \big|X_i^{(N)}\big| > \varepsilon\Big) \\ &\quad \le \text{constant} \times E[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N} \\ &\quad \to 0, \end{aligned} \tag{5.20}$$

Jacod and Shiryaev [(1987), Proposition VI.3.26, page 315] yields that any limit $\tilde{l}_t$ must be sample path continuous. As (5.14) yields that $\tilde{l}_t^{(N)2} - \langle \tilde{l}^{(N)}, \tilde{l}^{(N)}\rangle_t$ is a uniformly integrable family of random variables, Jacod and Shiryaev (1987), Proposition IX.1.12, page 484, gives that $\tilde{l}_t^2 - (t \wedge b^2)$ must be a martingale. It follows that $\tilde{l}_t$ is a Wiener process up to time b^2, and, thereafter, constant. Hence $\tilde{l}_t^{(N)}$ converges in law in $\mathbb{D}(\mathbb{R}^+)$ to such a process $\tilde{l}_t$.

As for $\tilde{m}_t^{(N)}$, on the other hand, we can only assert compactness in terms of convergence of finite-dimensional distributions. Let $\tilde{m}_t$ be a limit point under such convergence. By uniform integrability, $\tilde{m}_t$ is a martingale, and so is $\tilde{l}_t\tilde{m}_t$. Using the representation theorem for functionals of a Wiener process [see, e.g., Theorem 11.16 (page 347) of Jacod (1979)], it follows that $\tilde{m}_t$ is independent of $\tilde{l}_t$, whence $E(m|l) = E(\tilde{m}_{\bar{k}+1}|\tilde{l}_{\bar{k}+1}) = 0$ a.s. □

PROOF OF PROPOSITION 2. Assume (5.14). As in Theorem 1, this is without loss of generality. Using Theorem II.2 (page 273) of Rebolledo (1980), it is clear that the martingale $\tilde{\tilde{l}}_t^{(N)}$, given by

$$\tilde{\tilde{l}}_t^{(N)} = l_i^{(N)}/T_N \quad \text{for } \frac{i}{T_N} \le t < \frac{i+1}{T_N}, \tag{5.21}$$

converges weakly to a Wiener process $\tilde{\tilde{l}}_t$ in $\mathbb{D}([0,1])$ (with the Skorokhod topology).

On the other hand, set

$$\omega_t^{(N)} = \left[\tilde{\tilde{l}}^{(N)}, \tilde{\tilde{l}}^{(N)}\right]_t - \left\langle \tilde{\tilde{l}}^{(N)}, \tilde{\tilde{l}}^{(N)} \right\rangle_t, \tag{5.22}$$

and note that

$$\begin{aligned} E\left(\omega_1^{(N)}\right)^2 &\le E[l^{(N)}, l^{(N)}, l^{(N)}, l^{(N)}]_{T_N} \\ &= O\left(r_N^2\right), \end{aligned} \tag{5.23}$$

whence $(r_N^{-1}\omega_t^{(N)})_{0 \le t \le 1}$ is compact as far as convergence of finite-dimensional distributions (fidi) is concerned.

Furthermore, (5.14), (5.23) and the Jensen and Hölder inequalities yield that

$$E\left|\left\langle \tilde{\tilde{l}}^{(N)}, \omega^{(N)} \right\rangle_t\right| = O(r_N), \tag{5.24}$$

whence $(r_N^{-1}\langle \tilde{\tilde{l}}^{(N)}, \omega^{(N)}\rangle_t)_{0 \le t \le 1}$ is also compact in terms of fidi convergence. Since

$$\left\langle \tilde{\tilde{l}}^{(N)}, \omega^{(N)} \right\rangle_t = \langle l^{(N)}, l^{(N)}, l^{(N)} \rangle_{i/T_N} \quad \text{for } \frac{i}{T_N} \le t < \frac{i+1}{T_N}, \tag{5.25}$$

the assumption of the proposition yields that any fidi limit point of $r_N^{-1}\langle \tilde{\tilde{l}}^{(N)}, \omega^{(N)} \rangle_t$ must be nonrandom.

Hence, using the representation property for functionals of a Wiener process [see Theorem 11.16 (page 347) of Jacod (1979)], together with the fact that $\tilde{\tilde{l}}_t$ is a Wiener process, it follows that for any fidi limit ω_t of $\omega_t^{(N)}$, $E(\omega_1 | \tilde{\tilde{l}}_1)$ is a linear function of $\tilde{\tilde{l}}_1$. Since $\langle \tilde{\tilde{l}}, \omega \rangle_1 = \lim_{N \to \infty} r_N^{-1}\langle l^{(N)}, l^{(N)}, l^{(N)} \rangle_{T_N}$, the result follows. □

PROOF OF PROPOSITION 3. To indicate the dependence on N, we refer to quantities generated under Q_N, θ_N by superscript (N). $\|\cdot\|_{N,p}$ is the L^p-norm for this probability space.

Assume that $|\theta_N| \le 1 - k_0$, for some $k_0 > 0$ (this will eventually happen). Also assume (with no loss of generality) that we are dealing with a subsequence so that, as $N \to \infty$,

$$E^{(N)}\big(\varepsilon_1^{(N)}\big)^4 \to \kappa_4, \tag{5.26}$$

$$\sqrt{N}\left(E^{(N)}\big(\varepsilon_1^{(N)}\big)^2 - \kappa_2\right) \to \alpha, \tag{5.27}$$

$$\sqrt{N}\,(\theta_N - \theta) \to \alpha'. \tag{5.28}$$

Set

$$\begin{aligned} M_t^{(N)} &= \eta_t^{(N)^2} - \eta_0^{(N)^2} + \big(1 - \theta_N^2\big) \sum_{s=0}^{t-1} \eta_s^{(N)^2} - tE^{(N)}\big(\varepsilon^{(N)}\big)^2 \\ &= \sum_{s=0}^{t-1} \Big[\eta_{s+1}^{(N)^2} - \theta_N^2 \eta_s^{(N)} - E^{(N)}\big(\varepsilon^{(N)}\big)^2\Big] \\ &= \sum_{s=0}^{t-1} \Big[2\theta \eta_s^{(N)} \varepsilon_{s+1}^{(N)} + \varepsilon_{s+1}^{(N)^2} - E^{(N)}\big(\varepsilon^{(N)}\big)^2\Big]. \end{aligned} \tag{5.29}$$

This is a martingale, which will be useful below. Also, set, for $k = 2, 3$:

$$\rho_k = \kappa_k \big(1 - \theta^k\big)^{-1}. \tag{5.30}$$

From (5.29) and (5.30), we have that

$$\begin{aligned} \sqrt{N}\left(\frac{1}{N} \sum_{s=0}^{t-1} \eta_s^{(N)^2} - \rho_2\right) &= \big(1 - \theta_N^2\big)^{-1} N^{-1/2} M_N^{(N)} \\ &\quad - \big(1 - \theta_N^2\big)^{-1} N^{-1/2} \Big(\eta_t^{(N)^2} - \eta_0^{(N)^2}\Big) \\ &\quad + \big(1 - \theta_N^2\big)^{-1} N^{1/2} \Big(E^{(N)}\big(\varepsilon^{(N)}\big)^2 - \kappa_2\Big) \\ &\quad + N^{1/2}\Big[\big(1 - \theta_N^2\big)^{-1} - \big(1 - \theta^2\big)^{-1}\Big]\kappa_2. \end{aligned} \tag{5.31}$$

We will bound and find a limit for the l.h.s. with the help of this decomposition.

Begin by observing that since

$$E^{(N)} \eta_{s+1}^{(N)^2} = \theta_N^2 E^{(N)} \eta_s^{(N)^2} + E^{(N)}\big(\varepsilon^{(N)}\big)^2,$$

we have that

$$E^{(N)} \eta_t^{(N)^2} \le \theta_N^{2t} x_0^2 + \big(1 - \theta^2\big)^{-1} E^{(N)}\big(\varepsilon^{(N)}\big)^2. \tag{5.32}$$

Since x_0 is fixed and $|\theta_N| \le 1 - k_0$, we get that if $0 < \delta < 1$,

$$\begin{aligned}
&\left\| \sqrt{N}\left(\frac{1}{N}\sum_{s=0}^{N-1}\eta_s^{(N)^2} - \rho_2\right)\right\|_{N,1+\delta} \\
&\le k_1\left(\left\|\frac{M_N^{(N)}}{\sqrt{N}}\right\|_{N,1+\delta} + \left|N^{1/2}(\theta_N - \theta)\right| + k_2\right) \\
&\qquad\qquad \text{(where } k_1 \text{ and } k_2 \text{ are constants)} \qquad (5.33)\\
&\le k_3\left\|\left(\varepsilon^{(N)}\right)^2\right\|_{N,(1+\delta)/2}\frac{1}{N}\sum_{s=0}^{N-1}\left\|\eta_s^{(N)^2}\right\|_{N,(1+\delta)/2} \\
&\quad + k_3\left(\left\|\left(\varepsilon^{(N)}\right)^4\right\|_{N,(1+\delta)/2} + \left|N^{1/2}(\theta_N-\theta)\right| + k_2\right)
\end{aligned}$$

by Burkholder's inequality. By the "telescope" principle used in (5.29),

$$\begin{aligned}
&\left(1-\theta_N^2\right)\sum_{s=0}^{N-1}\left\|\eta_s^{(N)^2}\right\|_{N,(1+\delta)/2} \\
&\le \left\|\eta_N^{(N)^2}\right\|_{N,(1+\delta)/2} + x_0^2 \\
&\quad + 2|\theta_N|\,\left\|\varepsilon^{(N)}\right\|_{N,(1+\delta)/2}\sum_{s=0}^{N-1}\left\|\eta_s^{(N)}\right\|_{N,(1+\delta)/2} \qquad (5.34)\\
&\quad + N\left\|\left(\varepsilon^{(N)}\right)^2\right\|_{N,(1+\delta)/2} \\
&\le \left(1+\theta_N^{2N} + \left(1-\theta_N^2\right)^{-1}\right)x_0^2 \\
&\quad + \left[\left(1-\theta_N^2\right)^{-1} + \left(1 + \left(1-\theta_N^2\right)^{-1}\right)N\right]E^{(N)}\left(\varepsilon^{(N)}\right)^2,
\end{aligned}$$

by (5.32) and Liapunov's inequality. That inequality, together with (5.33) and (5.34), now yields that

$$\begin{aligned}
&\left\| \sqrt{N}\left(\frac{1}{N}\sum_{s=0}^{N-1}\eta_s^{(N)^2} - \rho_2\right)\right\|_{N,1+\delta} \qquad (5.35)\\
&\le k_4\left(E^{(N)}\left(\varepsilon^{(N)}\right)^4 + \left|N^{1/2}(\theta_N-\theta)\right| + k_5\right).
\end{aligned}$$

Hence the integrability condition on the square variation is satisfied, with $\underline{k} = 0$, $\overline{k} = \infty$.

If we set

$$D_N^* = \left\{\left|\frac{1}{N}\sum_{s=0}^{N-1}\eta_s^{(N)^2} - \rho_2\right| \ge k_1^* \text{ or } \left|\frac{1}{N}\sum_{s=0}^{N-1}\eta_s^{(N)}\varepsilon_{s+1}^{(N)}\right| \ge k_2^*\right\}, \qquad (5.36)$$

where $k_1^*, k_2^* > 0$, then the integrability condition on $\sigma_N = s_N(\sum_{s=0}^{N-1}\eta_s^{(N)^2})^2$ is also satisfied.

Furthermore,

$$E^{(N)}\eta_{s+1}^{(N)^4} = \theta_N^4 E^{(N)}\eta_s^{(N)^4} + 6\theta_N^2 E^{(N)}\eta_s^{(N)^2} E^{(N)}\left(\varepsilon^{(N)}\right)^2 + 4\theta_N x_0 E^{(N)}\left(\varepsilon^{(N)}\right)^3 + E^{(N)}\left(\varepsilon^{(N)}\right)^4, \tag{5.37}$$

whence, by a telescoping sum,

$$\begin{aligned}
\left(1-\theta_N^4\right)\sum_{s=0}^{N-1} E^{(N)}\eta_s^{(N)^4}
&= E^{(N)}\eta_N^{(N)^4} - x_0^4 + 6\theta_N^2 E^{(N)}\left(\varepsilon^{(N)}\right)^2 \sum_{s=0}^{N-1} E^{(N)}\eta_s^{(N)^2} \\
&\quad + N4\theta_N x_0 E^{(N)}\left(\varepsilon^{(N)}\right)^3 + NE^{(N)}\left(\varepsilon^{(N)}\right)^4 \\
&\le 12\theta_N^2 E^{(N)}\left(\varepsilon^{(N)}\right)^2 \sum_{s=0}^{N-1} E^{(N)}\eta_s^{(N)^2} \\
&\quad + N8\theta_N x_0 E^{(N)}\left(\varepsilon^{(N)}\right)^3 + N2E^{(N)}\left(\varepsilon^{(N)}\right)^4 \quad \text{[by (5.37) again]} \\
&\le N \,\text{constant}\left(E^{(N)}\left(\varepsilon^{(N)}\right)^4 + 1\right),
\end{aligned} \tag{5.38}$$

by (5.32) and since $|\theta_N| \le 1 - k_0$. Hence the integrability condition for the fourth-order variation is satisfied.

It remains to find the ψ's. Using the techniques from (5.29) (telescoping sum, representation as martingale) and also the uniform integrability,

$$\frac{1}{N}\sum_{s=0}^{N-1} \eta_s^{(N)^k} \to \rho_k \tag{5.39}$$

in probability, for $k = 2, 3, 4$, where, in particular, ρ_2 and ρ_3 are as given by (5.30). By Corollary 3.1 in Hall and Heyde (1980) and by (5.29), it then follows that

$$\left(\frac{1}{\sqrt{N}}\sum_{s=0}^{N-1}\eta_s^{(N)}\varepsilon_{s+1}^{(N)}, \frac{M_N^{(N)}}{\sqrt{N}}, \frac{1}{\sqrt{N}}\sum_{s=0}^{N-1}\left(\eta_s^{(N)}\right)^2\left(\left(\varepsilon_{s+1}^{(N)}\right)^2 - E\left(\varepsilon_{s+1}^{(N)^2}\right)\right)\right) \to \left(\sqrt{\kappa_2\rho_2}\,Z, U, V\right) \tag{5.40}$$

in law, where (Z, U, V) is trivariate normal with mean 0 and with

$$\operatorname{Cov}(Z, U) = 2\theta\sqrt{\kappa_2\rho_2}, \tag{5.41}$$

$$\operatorname{Cov}(Z, V) = (\kappa_2\rho_2)^{-1/2}\kappa_3\rho_3. \tag{5.42}$$

In view of (5.31), we can take

$$\psi_p(z) = 2\theta\left(1-\theta^2\right)^{-1/2} z + 2\kappa_2^{-1}\alpha + 2\theta\left(1-\theta^2\right)^{-1}\alpha', \tag{5.43}$$

$$\psi_o(z) - \psi_p(z) = (\kappa_2\rho_2)^{-3/2}\kappa_3\rho_3 z \tag{5.44}$$

and

$$\psi_*(z) = \psi_p(z). \tag{5.45}$$

This proves the result. □

PROOF OF PROPOSITION 5. Refer to the notation of Nummelin (1984). Assume that $m_0 = 1$; the proof is similar in the general case.

First note that, since σ_N/N can be taken to be bounded away from 0 on D_N^* without loss of generality (use Chebyshev's inequality),

$$\begin{aligned} &E_\lambda\Big(g\big((l_{N+T_\alpha} - l_{T_\alpha})/\sqrt{\sigma_N}\big) - g\big(l_N/\sqrt{\sigma_N}\big)\Big)I_{D_N^*}I_{(T\alpha<N)} \\ &\quad = N^{-1/2}b_*^{-1}E\Big((l_{N+T_\alpha} - l_{T_\alpha} - l_N) \\ &\qquad\qquad \times g'\big(l_N/b^*\sqrt{N}\big)I_{(T_\alpha<N)}\Big) + o(N^{-1/2}) \\ &\quad = o(N^{-1/2}), \end{aligned} \tag{5.46}$$

since $E_\lambda((l_{N+T_\alpha} - l_N)I_{(T_\alpha<N)}|\mathscr{F}_N) = 0$ and by combining Corollary 3.2 (page 64) in Hall and Heyde (1980) with Theorem 4.3.6 (page 123) in Revuz (1975). This is provided we can show that $E_\lambda|l_{N+T_\alpha} - l_{T_\alpha}|I_{(T_\alpha<N)}$ and $E_\lambda|l_{T_\alpha}|$ are bounded. The latter is finite in view of (4.8) and Proposition 5.13 (page 80) in Nummelin (1984). The former is finite because

$$\begin{aligned} &E_\lambda\left|\sum_{i=N+1}^{N+T_\alpha} f(\eta_i)\right| I(T_\alpha<N) \\ &\quad = \sum_{k=1}^{N-1} P_\lambda(T_\alpha = k)E_\lambda\left\{\left|\sum_{i=N+1}^{N+k} f(\eta_i)\right|\Bigg|T_\alpha = k\right\} \\ &\quad = \sum_{k=1}^{N-1} P_\lambda(T_\alpha = k)E_\nu\left|\sum_{N+1-k}^{N} f(\eta_i)\right| \\ &\quad \le \sum_{k=1}^{N-1} P_\lambda(T_\alpha = k)k \times \text{constant} \times E_\pi|f| \\ &\quad < \infty \end{aligned} \tag{5.47}$$

by (4.3), (4.5) and Nummelin (1984), Proposition 5.13.

Since $P_\lambda(T_\alpha \ge N) \le N^{-1}E_\lambda T_\alpha = o(N^{-1/2})$ for the same reason, it follows from (5.46) that it is enough to show the expansion for $(l_{N+T_\alpha} - l_{T_\alpha})/\sqrt{\sigma_N}$. That boils down to verifying the integrability and CLT conditions for l_N under the initial distribution ν.

To proceed, we need two slight modifications of the results in Nummelin (1984).

LEMMA 3. *Theorem* 6.12, *first part, of Nummelin* (1984) *holds provided* (η_N) *is ergodic*, $E_\pi|f(\eta)| < \infty$, *and the initial measure and* π *are* 1- *and* $|f|$-*regular*.

The proof of this is an obvious extension of Nummelin's proof.

LEMMA 4. *Let h be a function satisfying the conditions of Nummelin's Theorem* 7.6. *Then*

$$N^{-1/2}\sum_{i=0}^{N} h(\eta_t) \xrightarrow{\mathscr{L}} N(0, s_h^2). \tag{5.48}$$

The sequence is uniformly integrable under π, and $s_h^2 = \pi h^2 + 2\pi hP\overline{G}_{m_0, s, \nu}h$.

PROOF. We shall show that

$$E_\pi\left[N^{-1/2}\sum_{i=0}^{N} h(\eta_i)\right]^2 \to s_h^2. \tag{5.49}$$

One can similarly show that

$$E_\pi\left[(Nm_0)^{-1/2}\sum_{i=0}^{Nm_0} h(\eta_i)\right]^2 \to \sigma_h^2, \tag{5.50}$$

where m_0 and σ_h^2 refer to Nummelin's notation. Clearly, $s_h^2 = \sigma_h^2$. Since, furthermore, Nummelin's Theorem 7.6 states that $N^{-1/2}\sum_{i=1}^{N} h(\eta_i)$ is asymptotically normal with mean 0 and variance σ_h^2, our Lemma 4 holds.

It remains to show (5.49). Write

$$E_\pi\left[N^{-1/2}\sum_{i=0}^{N} h(\eta_i)\right]^2 = \pi f^2 + 2\frac{1}{N}\sum_{v=0}^{N} u_v, \tag{5.51}$$

where $u_v = \pi f \sum_{i=1}^{v} P^i f$. Equation (5.49) now follows from Lemma 3 and the Toeplitz lemma [see, e.g., Hall and Heyde (1980), page 31]. □

Since $v \ll \pi$ and $dv/d\pi$ is essentially bounded, the integrability condition on $[l, l, l, l]_N$ follows automatically from our assumptions. For the same reason, the integrability condition on $[l, l]_N$ follows from Lemma 4. The existence and form of ψ_o follows from setting $h = af + b(f^2 - \pi f^2)$ in Lemma 2, and using the Cramér–Wold device. The existence and form of ψ_p then follows from Proposition 2 and Theorem 4.3.6 (page 123) of Revuz (1975). □

Acknowledgments. The author would like to thank Stephen Stigler, Wing Wong and a referee for very helpful discussions and comments, and Mitzi Nakatsuka for typing the manuscript.

REFERENCES

ABRAMOVITCH, L. and SINGH, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.* **13** 116–132.

ANDERSEN, P. K. and BORGAN, Ø. (1985). Counting process models for life history data: A review (with discussion). *Scand. J. Statist.* **12** 97–158.

ANDERSEN, P. K. and GILL, R. D. (1982). Cox's regression model for counting processes: A large sample study. *Ann. Statist.* **10** 1100–1120.

BARNDORFF-NIELSEN, O. E. and COX, D. R. (1984). Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator. *J. Roy. Statist. Soc. Ser. B* **46** 483–495.

BERAN, R. (1987). Prepivoting to reduce level error of confidence sets. *Biometrika* **74** 457–468.

BERAN, R. (1988). Prepivoting test statistics. *J. Amer. Statist. Assoc.* **83** 687–697.

BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451.

BHATTACHARYA, R. N. and RAO, R. R. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

BHATTACHARYA, R. N. and RAO, R. R. (1986). *Normal Approximation and Asymptotic Expansions.* Krieger, Malabar, FL.

BOLTHAUSEN, E. (1982). Exact convergence rates in some martingale central limit theorems. *Ann. Probab.* **10** 672–688.

BOSE, A. (1986). Asymptotic study of estimators in some discrete and continuous time models. Ph.D. dissertation, Indian Statist. Inst., Calcutta.

BOSE, A. (1987). Bootstrapping in moving average models. Technical Report 87-55, Purdue Univ.

BOSE, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Ann. Statist.* **16** 1709–1722.

BRÉMAUD, P. (1981). *Point Processes and Queues—Martingale Dynamics.* Springer, New York.

CHAN, K.-S. (1990). Deterministic stability, stochastic stability, and ergodicity. In *Non-Linear Times Series. A Dynamical System Approach* (H. Tong, ed.) 448–466. Oxford Univ. Press.

CORNISH, E. A. and FISHER, R. A. (1937). Moments and cumulants in the specification of distributions. *Rev. Inst. Internat. Statist.* **4** 1–14.

CRAMÉR, H. (1937). *Random Variables and Probability Distributions. Cambridge Tracts in Math.* **36**. Cambridge Univ. Press.

DAMBIS, K. (1965). On the decomposition of continuous sub-martingales. *Theory Probab. Appl.* **10** 401–410.

DUBINS, L. E. and SCHWARTZ, G. (1965). On continuous martingales. *Proc. Nat. Acad. Sci. U.S.A.* **53** 913–916.

EDGEWORTH, F. Y. (1883). The law of error. *Philosophical Mag.* **16** 300–309.

EDGEWORTH, F. Y. (1905). The law of error. *Cambridge Philos. Trans.* **20** 36–65, 113–141.

EFRON, B. (1987). Better bootstrap confidence intervals. *J. Amer. Statist. Assoc.* **82** 171–185.

GOETZE, F. and HIPP, C. (1978). Asymptotic expansions in the central limit theorm under moment conditions. *Z. Wahrsch. Verw. Gebiete* **42** 67–87.

GOETZE, F. and HIPP, C. (1983). Asymptotic expansions for sums of weakly dependent random vectors. *Z. Wahrsch. Verw. Gebiete* **64** 211–239.

HAEUSLER, E. (1988). On the rate of convergence in the central limit theorem for martingales with discrete and continuous time. *Ann. Probab.* **16** 275–299.

HALL, P. (1986). On the bootstrap and confidence intervals. *Ann. Statist.* **14** 1431–1452.

HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals. *Ann. Statist.* **16** 927–953.

HALL, P. and HEYDE, C. C. (1980). *Martingale Limit Theory and Its Application.* Academic, New York.

HEATH, D. (1977). Interpolation of martingales. *Ann. Probab.* **5** 804–806.

JACOD, J. (1979). *Calcul Stochastique et Problemes de Martingales. Lecture Notes in Math.* **714**. Springer, Berlin.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes.* Springer, Berlin.

JENSEN, J. L. (1986). Asymptotic expansions for sums of dependent variables. Memoir 10, Dept. Theoret. Statist., Univ. Aarhus.

JENSEN, J. L. (1989). Asymptotic expansions for strongly mixing Harris recurrent Markov chains. *Scand. J. Statist.* **16** 47–64.

KLIMKO, L. A. and NELSON, P. I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* **6** 629–642.

MALINOVSKII, V. K. (1987). Limit theorems for Harris Markov chains. *Theory Probab. Appl.* **31** 269–285.

MCCULLAGH, P. (1984a). Local sufficiency. *Biometrika* **71** 233–244.

MCCULLAGH, P. (1984b). Tensor notation and cumulants of polynomials. *Biometrika* **71** 461–476.

MYKLAND, P. A. (1989). Bootstrap and Edgeworth methods for dependent variables. Ph.D. dissertation, Univ. California, Berkeley.

MYKLAND, P. A. (1992). Asymptotic expansions and bootstrapping-distributions for dependent variables: A martingale approach. *Ann. Statist.* **20** 623–654.

NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators.* Cambridge Univ. Press.

REBOLLEDO, R. (1980). Central limit theorems for local Martingales. *Z. Wahrsch. Verw. Gebiete* **51** 269–286.

REVUZ, D. (1975). *Markov Chains.* North-Holland, Amsterdam.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

TANAKA, K. (1986). Asymptotic expansions for time series statistics. *J. Appl. Probab.* **23A** 211–227.

TANIGUCHI, M. (1984). Validity of Edgeworth expansions for statistics of time series. *J. Time Ser. Anal.* **5** 37–52.

TJØSTHEIM, D. (1986). Estimation in nonlinear time series models. *Stochastic Process Appl.* **21** 251–273.

TONG, H. (1990). *Non-Linear Time Series. A Dynamical System Approach.* Oxford Univ. Press.

WONG, W. H. (1986). Theory of partial likelihood. *Ann. Statist.* **14** 88–123.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
5734 UNIVERSITY AVENUE
CHICAGO, ILLINOIS 60637